we describe a practical and scalable algorithm for determining candidate network interactions based on decomposing an *N*-dimensional system into *N* one-dimensional problems. The algorithm was tested on *in silico* networks based on known biological GRNs. The computational complexity of the network identification is shown to increase as N^2 while a parallel implementation achieves essentially linear speedup with the increasing number of processing cores. For each *in silico* network tested, the algorithm successfully predicts a candidate network that reproduces the network dynamics. This approach dramatically reduces the computational demand required for reverse engineering GRNs and produces a wealth of exploitable information in the process. Moreover, the candidate network topologies returned by the algorithm can be used to design future experiments aimed at gathering informative data capable of further resolving the true network topology.

919-Pos Board B705
Multiple Co-Evolutionary Networks have Evolved on the Common Tertiary Scaffold of the LacI/GalR Proteins
Liskin Swint-Kruse, Daniel J. Parente.
KU Medical Center, Kansas City, KS, USA.
The number of known protein sequences is growing at a phenomenal rate. A first step in understanding the functional implications of novel sequences is to extract information from characterized, homologous proteins via sequence/function analyses. To that end, many new algorithms have been recently developed. Of these, co-evolutionary analysis has emerged as a popular class of methods. These methods identify pairs of positions in a multiple sequence alignment that vary in a coordinated manner. From these calculations, a network of co-evolving positions can be assembled. Strongly coevolving positions are thought to participate in the same functional and/or structural role.
Generally, co-evolution analyses are carried out on protein families with >40% sequence identity. One unresolved question is the extent to which co-evolutionary networks are conserved on a common tertiary architecture. To investigate this question, we are comparing the co-evolutionary networks of various subfamilies in the LacI/GalR transcription repressor family. Each subfamily has >40% sequence identity. In contrast, identity between subfamilies is usually 15-25%. Nevertheless, available structures show that subfamilies have very similar tertiary structures. Co-evolution networks have been generated using five different algorithms. As previously reported, the algorithm outputs differ from each other (e.g., McBASC and OMES identify different co-evolving positions); however, all algorithms find different patterns for each LacI/GalR subfamily. Additional analyses show that (i) each subfamily has a different set of conserved positions; and (ii) positions that are conserved in one subfamily are highly varied in another. Thus, the common tertiary architecture of the LacI/GalR homologs accommodates a variety of functional and/or structural networks that is distinct for divergent subfamilies.

920-Pos Board B706
An Improved Meta-Analysis Procedure Reveals Key Transcriptome Signatures Underlying a Renal Damage Phenotype
Yu Bai, Jennifer Schmahl.
Regeneron, Tarrytown, NY, USA.
Integration of transcriptome profiling data via meta-analysis can significantly leverage statistic power while reducing false positives and thus is widely applied in genomic research. Various meta-analysis procedures, ranging from Venn diagram to score-based gene selections, have been employed to uncover key genes underlying a phenotype or disease. Scores are often adopted to characterize genes beyond the simple knowledge of being perturbed or not and permit statistical control of false positives. Nevertheless, many scoring methods use significance (e.g. p-value, q-value) to characterize the gene perturbation instead of more biology relevant measures such as fold change. Recent work showed the advantage of fold change-ranked scores. However, the rank of genes can vary significantly for small alterations in fold change and therefore noise-sensitive. Herein we developed a new score combining fold change and the occurrence counts to select genes that are most commonly and strongly perturbed, with false discovery rate (FDR) estimated by permutation. The method was evaluated in comparison to previous rank or occurrence scores using synthetic data, where multiple datasets were simulated with predefined meta-genes (noise introduced so not all meta-genes are perturbed in each dataset) in the background of a varying number of other randomly perturbed genes. Our method outcompeted previous occurrence score with higher sensitivity, and the rank score with both higher sensitivity and specificity. We applied the method to dissect a renal damage phenotype observed in a set of KO mice with unknown mechanism, yielding 247 meta-genes (FDR<=0.1) enriched in renal and urological function (pvalue<1e-4, Ingenuity) and significantly correlated with a Cisplatin-induced acute renal failure model (pvalue<1e-35, NextBio). These insightful meta-signatures were not achieved by other methods, highlighting the power of our new method in facilitating deeper understandings of the molecular mechanism underlying disease phenotypes.

921-Pos Board B707
In Silico Reconstruction of HIV Viral Fitness Landscapes
Andrew L. Ferguson, Arup K. Chakraborty.
MIT, Cambridge, MA, USA.
Since its discovery in the early 1980's, HIV/AIDS has claimed over 30 million lives worldwide, with another 33.3 million infected today and no effective vaccine yet available. Contemporary vaccine design efforts adopt a new paradigm to counter the extreme mutability of HIV by targeting vulnerable regions of the proteome in which mutations to evade the immune response fatally compromise viral fitness. In this way HIV is trapped between a rock - vaccine induced immune pressure - and a hard place - crippled mutant strains. We have developed maximum entropy models, which map to an Ising spin glass Hamiltonian, fitted to multiple sequence alignment data, to permit the *in silico* construction of sequence space fitness landscapes of HIV. These quantitative models of viral fitness allow identification of T-cell epitopes that maximally abolish fitness and block viral escape mutations, providing a systematic means to identify promising candidate regions for next generation vaccine trials. Clinical validation of our approach is provided by its identification of vulnerable regions in the *gag* polyprotein that are naturally targeted by rare patients who control HIV infection without therapy, and reproduction within our fitness landscapes of well-documented viral escape pathways.

922-Pos Board B708
How to Find Non Hub Important Nodes in Protein Networks?
Vince Grolmusz, Gabor Ivan, Daniel Banky, Balazs Szerencsi.
Eötvös University, Budapest, Hungary.
Biological network data, such as metabolic-, signaling- or physical interaction graphs of proteins are increasingly available in public repositories for important species. Tools for the quantitative analysis of these networks are being developed today. Protein network-based drug target identification methods usually returns protein hubs with large degrees (also called "connectivity") in the networks as potentially important targets. Some known, important protein targets, however, are not hubs at all, and on the other hand, perturbing protein hubs in these networks may have several unwanted physiological effects, due to their interaction with numerous partners.
Here we show a novel method applicable in networks with directed edges (such as metabolic networks) that compensates the low degree (non-hub) vertices in the network, and identifies intrinsically important nodes, totally independently from their hub properties. We demonstrate that the method correctly finds numerous already validated drug targets in distinct organisms (Mycobacterium tuberculosis, Plasmodium falciparum and MRSA Staphylococcus aureus), and consequently, it may suggest new possible protein targets as well.
Our method computes the PageRank for the nodes of the network, and divides the PageRank by the in-degree (i.e., the number of incoming edges) of the node. This quotient is the same in all nodes in an undirected graph (even for large- and low degree nodes, that is, for hubs and non-hubs as well), but may differ significantly from node to node in directed graphs. We suggest to assign importance to non-hub nodes with large PageRank/in-degree quotient.

923-Pos Board B709
A Multiscale View of Protein-Protein Interactions
Franca Fraternali.
King's College, London, United Kingdom.
High-throughput proteomics has allowed for the drafting of large Protein-Interaction Networks and systems biology devotes a constant effort in the statistical analysis of these. PPIN derived from large-scale experiments portrait a global picture of proteins connectivity and offer an intricate and complex picture of the extracted interactions. The detailed analysis of a particular cellular function may benefit from the study of smaller and accurately selected sub-networks, containing high confidence interactions resulting from the cross-mapping of multiple sources of information for the nodes such as tissue-specific expression data, gene expression data, domain profiles and structural information. We call these extracted networks Sub-Networks Knowledge. Analysis of these SNKs provides the basis for the identification of fundamental components of the network and of specific pathways associated with the studied phenomenon that can be targeted to speed-up experimental screening and design new experiments.
By reducing further the scale of details analysed in a given PPIN, one could focus only on special proteins in the network, like the multi-partner ones, called 'hubs'. In particular, we concentrated on protein hubs that have a 3D-structure and on the characterization of their conformational dynamics and flexibility of the multiple interfaces. A large-scale survey of the intrinsic dynamics of protein interfaces has been done using a dataset of 340 proteins classified according to the different number of partners. Conformational ensembles were generated